



用于网络文件系统 (NFS) 的 Sun Cluster 数据服务指南 (适用于 Solaris OS)

Sun Microsystems, Inc.
4150 Network Circle
Santa Clara, CA 95054
U.S.A.

文件号码: 819-1245-10
2004 年 9 月, 修订版 A

版权所有 2004 Sun Microsystems, Inc. 4150 Network Circle, Santa Clara, CA 95054 U.S.A. 保留所有权利。

本产品或文档受版权保护，并按照限制其使用、复制、发行和反汇编的许可证进行分发。未经 Sun 及其许可方的事先书面许可，不得以任何形式、任何手段复制本产品或文档的任何部分。第三方软件，包括字体技术，均已从 Sun 供应商处获得版权和使用许可。

本产品的某些部分可能是从 Berkeley BSD 系统衍生出来的，并获得了加利福尼亚大学的许可。UNIX 是 X/Open Company, Ltd. 在美国和其他国家/地区独家许可的注册商标。

Sun、Sun Microsystems、Sun 徽标、docs.sun.com、AnswerBook、AnswerBook2 和 Solaris 是 Sun Microsystems, Inc. 在美国和其他国家/地区的商标或注册商标。所有 SPARC 商标的使用均已获得许可，它们是 SPARC International, Inc. 在美国和其他国家/地区的商标或注册商标。标有 SPARC 商标的产品均基于由 Sun Microsystems, Inc. 开发的体系结构。

OPEN LOOK 和 Sun™ 图形用户界面是 Sun Microsystems, Inc. 为其用户和许可证持有者开发的。Sun 感谢 Xerox 在研究和开发可视或图形用户界面的概念方面为计算机行业所做的开拓性贡献。Sun 已从 Xerox 获得了对 Xerox 图形用户界面的非独占性许可证，该许可证还适用于实现 OPEN LOOK GUI 和在其他方面遵守 Sun 书面许可协议的 Sun 许可证持有者。

本文档按“原样”提供，对所有明示或默示的条件、陈述和担保，包括对适销性、适用性和非侵权性的默示保证，均不承担任何责任，除非此免责声明的适用范围在法律上无效。



041221@10536



目录

序 5

安装和配置 Sun Cluster HA for Network File System (NFS) 11

安装和配置 Sun Cluster HA for NFS 12

安装 Sun Cluster HA for NFS 软件包 12

 ▼ 如何使用 Web Start 程序安装 Sun Cluster HA for NFS 软件包 13

 ▼ 如何使用 `scinstall` 公用程序安装 Sun Cluster HA for NFS 软件包 14

注册和配置 Sun Cluster HA for NFS 14

 ▼ 如何注册和配置 Sun Cluster HA for NFS 15

 ▼ 如何更改 NFS 文件系统上的共享选项 19

 ▼ 如何动态更新 NFS 文件系统上的共享路径 21

 如何调整 Sun Cluster HA for NFS 方法超时 21

 如何配置 `SUNW.HAStoragePlus` 资源类型 22

配置 Sun Cluster HA for NFS 扩展特性 22

Sun Cluster HA for NFS 故障监视器 25

 启动故障监视器 25

 停止故障监视器 25

 NFS 系统故障监视器进程 25

 NFS 资源故障监视器进程 26

索引 29

序

《用于网络文件系统 (NFS) 的 Sun Cluster 数据服务指南 (适用于 Solaris OS)》说明了如何在基于 SPARC® 和基于 x86 的系统上安装和配置 Sun™ Cluster HA for Network File System (NFS)。

注意 – 在本文档中，术语“x86”是指 Intel 32 位微处理器芯片系列和 AMD 制造的兼容微处理器芯片。

此文档适用于具有丰富的 Sun 软硬件知识的系统管理员。不要将此文档作为规划指南或售前指南。在阅读此文档之前，您应该已经确定了系统要求并购买了相应的设备和软件。

此文档中的说明均假定读者了解 Solaris™ 操作系统，并熟练掌握与 Sun Cluster 软件一起使用的卷管理器软件。

注意 – Sun Cluster 软件可以在 SPARC 和 x86 两种平台上运行。除非在特定的章、节、说明、标有项目符号的项、图、表或示例中指出，否则本文档中的信息均适用于两种平台。

UNIX 命令

此文档包含针对于安装和配置 Sun Cluster 数据服务的特定命令的信息，但不包含基本 UNIX® 命令和过程（例如关闭系统、引导系统和配置设备）的完整信息。有关基本 UNIX 命令和过程的信息，可从以下来源获得：

- Solaris 操作系统的联机文档
- Solaris 操作系统手册页

- 系统附带的其它软件文档

印刷约定

下表说明了本书中使用的印刷约定。

表 P-1 印刷约定

字体或符号	含义	示例
AaBbCc123	命令、文件和目录的名称以及计算机屏幕输出	编辑 <code>.login</code> 文件。 使用 <code>ls -a</code> 列出所有文件。 <code>machine_name% you have mail.</code>
AaBbCc123	用户键入的内容，与计算机屏幕输出相对应	<code>machine_name% su</code> 口令：
<i>AaBbCc123</i>	命令行占位符：需用实名或值替换	删除文件的命令为 <code>rm filename</code> 。
<i>AaBbCc123</i>	书名、新增术语和要强调的术语	请参阅《 用户指南 》第 6 章。 这些称为 类选项 。 请勿保存文件。 (在联机状态下，强调有时显示为 粗体 。)

命令实例中的 shell 提示符

下表显示了 C shell、Bourne shell 和 Korn shell 的缺省系统提示符和超级用户提示符。

表 P-2 shell 提示符

Shell	提示符
C shell 提示符	<code>machine_name%</code>
C shell 超级用户提示符	<code>machine_name#</code>

表 P-2 shell 提示符 (续)

Shell	提示符
Bourne shell 和 Korn shell 提示符	\$
Bourne shell 和 Korn shell 超级用户提示符	#

相关文档

有关相关 Sun Cluster 主题的信息，可从下表列出的文档中获得。所有 Sun Cluster 文档均可从 <http://docs.sun.com> 获得。

主题	文档
数据服务管理	《Sun Cluster 数据服务规划和管理指南 (适用于 Solaris OS)》 单个数据服务指南
概念	《Sun Cluster 概念指南 (适用于 Solaris OS)》
概述	《Sun Cluster 概述 (适用于 Solaris OS)》
软件安装	《Sun Cluster 软件安装指南 (适用于 Solaris OS)》
系统管理	《Sun Cluster 系统管理指南 (适用于 Solaris OS)》
硬件管理	Sun Cluster 3.x Hardware Administration Manual for Solaris OS 单个硬件管理指南
数据服务开发	《Sun Cluster 数据服务开发者指南 (适用于 Solaris OS)》
错误消息	Sun Cluster Error Messages Guide for Solaris OS
命令和功能参考	Sun Cluster Reference Manual for Solaris OS

有关 Sun Cluster 文档的完整列表，请参阅 <http://docs.sun.com> 中的您的 Sun Cluster 版本的发行说明。

相关的第三方 Web 站点引用

本文档中引用的第三方 URL 提供了其他相关信息。

注意 – Sun 对本文档中述及的第三方 Web 站点的可用性不承担责任。Sun 对此类站点或资源上提供的或通过其得到的任何内容、广告、产品或其它材料不承担担保，也不承担任何责任或义务。Sun 不对由于使用或依赖于任何此类站点或资源上提供的或通过其得到的任何此类内容、商品或服务引起或宣称由之引起或与之有关的任何损害或损失负责。

联机访问 Sun 文档

可以通过 docs.sun.comSM Web 站点联机访问 Sun 技术文档。您可以浏览 docs.sun.com 档案或查找某个具体的书的标题或主题。URL 为 <http://docs.sun.com>。

订购 Sun 文档

Sun Microsystems 提供一些印刷的产品文档。有关文档列表以及如何订购它们，请参见 <http://docs.sun.com> 上的“购买印刷的文档”。

帮助

如果您在安装或使用 Sun Cluster 时有任何问题，请与您的服务供应商联系并提供以下信息：

- 您的姓名和电子邮件地址（如果有）
- 您的公司名称、地址和电话号码
- 系统的型号和序列号
- Solaris 操作系统的发行版本号（例如，Solaris 8）
- Sun Cluster 的发行版本号（例如，Sun Cluster 3.0）

使用以下命令可以针对为您的服务供应商收集系统中每个节点的信息。

命令	功能
<code>prtconf -v</code>	显示系统内存的大小并报告有关外围设备的信息
<code>psrinfo -v</code>	显示有关处理器的信息
<code>showrev -p</code>	报告已安装了哪些修补程序
<code>SPARC: prtdiag -v</code>	显示系统诊断信息
<code>scinstall -pv</code>	显示 Sun Cluster 发行版本和软件包版本信息

还请提供 `/var/adm/messages` 文件的内容。

安装和配置 Sun Cluster HA for Network File System (NFS)

本章说明了在 Sun Cluster 节点上安装和配置 Sun Cluster HA for Network File System (NFS) 的步骤。

本章包含以下过程。

- 第 13 页 “如何使用 Web Start 程序安装 Sun Cluster HA for NFS 软件包”
- 第 14 页 “如何使用 scinstall 公用程序安装 Sun Cluster HA for NFS 软件包”
- 第 15 页 “如何注册和配置 Sun Cluster HA for NFS”
- 第 19 页 “如何更改 NFS 文件系统上的共享选项”
- 第 21 页 “如何动态更新 NFS 文件系统上的共享路径”
- 第 21 页 “如何调整 Sun Cluster HA for NFS 方法超时”
- 第 22 页 “如何配置 SUNW.HAStoragePlus 资源类型”

您必须将 Sun Cluster HA for NFS 配置为故障切换数据服务。有关数据服务、资源组、资源和其它相关主题的一般信息，请参阅《*Sun Cluster 数据服务规划和管理指南（适用于 Solaris OS）*》中的“Planning for Sun Cluster Data Services”和《*Sun Cluster Concepts Guide for Solaris OS*》文档。

注意 – 您可以使用 SunPlex Manager 安装和配置此数据服务。有关详细信息，请参见 SunPlex Manager 联机帮助。

在安装和配置 Sun Cluster HA for NFS 之前，请使用《*Sun Cluster 数据服务规划和管理指南（适用于 Solaris OS）*》中的“Configuration Worksheets”中的工作表来规划您的资源和资源组。

在数据服务控制下的 NFS 装载点对于所有节点（可控制包含该文件系统的磁盘设备组）必须相同。

Sun Cluster HA for NFS 要求所有的 NFS 客户机安装都必须是“硬”安装。

Sun Cluster 节点不能是由 Sun Cluster HA for NFS 导出且受同一群集中某个节点控制的文件系统的 NFS 客户机。禁止以这种方式交叉安装 Sun Cluster HA for NFS。可以使用群集文件系统来在各群集节点间共享文件。

从 Solaris 9 开始，如果使用 Solaris 资源管理器管理分配给群集上的 NFS 的系统资源，则所有可以故障转移到普通群集节点的 Sun Cluster HA for NFS 资源必须具有相同的 Solaris 资源管理器项目 ID。项目 ID 由 Resource_project_name 资源特性设置。



注意 – 如果使用 VERITAS Volume Manager（仅在基于 SPARC 的群集中可以使用），则可以避免在 NFS 故障转移期间在客户机上出现“无效文件句柄”错误。请确保 vxio 驱动程序在所有群集节点上具有相同的伪设备主编号。完成安装后，您可以在 /etc/name_to_major 文件中查找此编号。

安装和配置 Sun Cluster HA for NFS

下表列出了介绍安装和配置任务的各个章节。

表 1 任务对应关系：安装和配置 Sun Cluster HA for NFS

任务	参考
安装 Sun Cluster HA for NFS 软件包	第 12 页 “安装 Sun Cluster HA for NFS 软件包”
设置和配置 Sun Cluster HA for NFS	第 14 页 “注册和配置 Sun Cluster HA for NFS”
配置资源扩展特性	第 22 页 “配置 Sun Cluster HA for NFS 扩展特性”
查看故障监视器信息	第 25 页 “Sun Cluster HA for NFS 故障监视器”

安装 Sun Cluster HA for NFS 软件包

如果未在 Sun Cluster 初始安装期间安装 Sun Cluster HA for Network File System (NFS) 软件包，请执行此过程以安装该软件包。在要安装 Sun Cluster HA for Network File System (NFS) 软件包的每个群集节点上执行此过程。必需使用 Sun Java Enterprise System Accessory CD Volume 3 才能完成此过程。

如果要同时安装多个数据服务，请执行《Sun Cluster 软件安装指南（适用于 Solaris OS）》的“安装软件”中的过程。

请使用以下安装工具之一来安装 Sun Cluster HA for Network File System (NFS) 软件包：

- Web Start 程序
- scinstall 公用程序

注意 – 在早于 Sun Cluster 3.1 数据服务 10/03 的发行版中未提供 Web Start 程序。

▼ 如何使用 Web Start 程序安装 Sun Cluster HA for NFS 软件包

您可以使用命令行界面 (CLI) 或图形用户界面 (GUI) 运行 Web Start 程序。CLI 和 GUI 中指令的内容和顺序类似。有关 Web Start 程序的详细信息，请参见 `installer(1M)` 手册页。

1. 在要安装 Sun Cluster HA for Network File System (NFS) 软件包的群集节点上成为超级用户。
2. (可选的) 如果要使用 GUI 运行 Web Start 程序，请确保已设置 `DISPLAY` 环境变量。
3. 将 Sun Java Enterprise System Accessory CD Volume 3 放入 CD-ROM 驱动器。如果卷管理守护程序 `vold(1M)` 正在运行并被配置为管理 CD-ROM 设备，它将在 `/cdrom/cdrom0` 目录中自动装载 CD-ROM。
4. 更改到 CD-ROM 的 Sun Cluster HA for Network File System (NFS) 组件目录。用于 Sun Cluster HA for Network File System (NFS) 数据服务的 Web Start 程序位于此目录中。

```
# cd /cdrom/cdrom0/  
components/SunCluster_HA_NFS_3.1
```

5. 启动 Web Start 程序。

```
# ./installer
```
6. 系统提示时，选择安装类型。
 - 要仅安装 C 语言环境，请选择“典型”。
 - 要安装其它语言环境，请选择“定制”。
7. 按照屏幕上的说明在节点上安装 Sun Cluster HA for Network File System (NFS) 软件包。

安装完成后，Web Start 程序将提供安装摘要。此摘要使您可以查看 Web Start 程序在安装期间创建的日志。这些日志位于 `/var/sadm/install/logs` 目录中。
8. 退出 Web Start 程序。
9. 从 CD-ROM 驱动器中取出 Sun Java Enterprise System Accessory CD Volume 3。
 - a. 为确保没有在使用 CD-ROM，请将目录更改为非 CD-ROM 中的目录。

b. 弹出 CD-ROM。

```
# eject cdrom
```

从此处可转到何处

转到第 14 页 “注册和配置 Sun Cluster HA for NFS”。

▼ 如何使用 `scinstall` 公用程序安装 Sun Cluster HA for NFS 软件包

本节介绍了如何使用 `scinstall` 公用程序安装 Sun Cluster HA for NFS 软件包。

1. 将 Sun Java Enterprise System Accessory CD Volume 3 放入 CD-ROM 驱动器。
2. 运行 `scinstall` 公用程序（不使用任何选项）。
此步骤将以交互模式启动 `scinstall` 公用程序。
3. 选择菜单选项“将新数据服务的支持添加到此群集节点”。
`scinstall` 公用程序将提示您输入其它信息。
4. 提供 Sun Java Enterprise System Accessory CD Volume 3 的路径。
公用程序将此 CD 称为“数据服务 CD”。
5. 指定要安装的数据服务。
`scinstall` 公用程序将列出您选定的数据服务，并要求您确认选择。
6. 退出 `scinstall` 公用程序。
7. 从驱动器中取出 CD。

从此处可转到何处

转到第 14 页 “注册和配置 Sun Cluster HA for NFS”。

注册和配置 Sun Cluster HA for NFS

此过程介绍了如何使用 `scrgadm(1M)` 命令注册和配置 Sun Cluster HA for NFS。

注意 – 通过其它选项也可以注册和配置数据服务。有关这些选项的详细信息，请参阅《*Sun Cluster 数据服务和规划指南（适用于 Solaris OS）*》中的“Tools for Data Service Resource Administration”。

注册和配置 Sun Cluster HA for NFS 之前，请运行以下命令以检验是否已在群集上安装了 Sun Cluster HA for NFS 软件包 SUNWscnfs。

```
# pkginfo -l SUNWscnfs
```

如果尚未安装该软件包，请参见第 12 页“安装 Sun Cluster HA for NFS 软件包”以获得有关如何安装该软件包的说明。

▼ 如何注册和配置 Sun Cluster HA for NFS

本节介绍了如何注册和配置 Sun Cluster HA for NFS。

1. 成为某个群集成员的超级用户。
2. 检验所有群集节点是否都处于联机状态。

```
# scstat -n
```

3. 创建 Pathprefix 目录。

Pathprefix 目录位于 Sun Cluster HA for NFS 用来维护管理和状态信息的群集文件系统中。

您可以指定任意目录用于此用途。但是，您必须为所创建的每个资源组手动创建一个 *admin-dir* 目录。例如，创建目录 */global/nfs*。

```
# mkdir -p /global/admin-dir
```

4. 创建故障切换资源组以包含 NFS 资源。

```
# scrgadm -a -g resource-group -y Pathprefix=/global/admin-dir [-h nodelist]
```

-a 指定您要添加新配置。

-g *resource-group* 指定故障切换资源组。

-y *Pathprefix =path* 在群集文件系统中指定 Sun Cluster HA for NFS 用来维护管理和状态信息的目录。

[-h *nodelist*] 指定一个可选的、用逗号分隔的物理节点名称或 ID 的列表，这些节点名称或 ID 标识潜在的主节点。列表中的顺序确定了 Resource Group Manager (RGM) 在故障切换过程中考虑主节点的顺序。

5. 检验是否已将您的所有逻辑主机名资源添加到名称服务数据库中。

为避免因名称服务查找而出现任何故障，请检验 Sun Cluster HA for NFS 使用的所有 IP 地址与主机名之间的映射是否都存在于服务器和客户机的 */etc/inet/hosts*

文件中。此要求同样适用于 Sun Cluster HA for NFS 使用的逻辑主机名的任何 IPMP 测试 IP 地址。

6. 在群集节点上的 `/etc/nsswitch.conf` 文件中配置名称服务映射，以在尝试访问 NIS 或 NIS+ 进行 `rpc` 查找之前先检查本地文件。

执行此操作可防止在公共网络或名称服务不可用时发生与计时相关的错误。

7. 修改 `/etc/nsswitch.conf` 中的主机项，以便在本地解析名称时，主机先不联系 NIS/DNS，而是立即返回一个成功状态。

执行此操作可使 HA-NFS 在发生公共网络故障时正确地进行故障切换。

```
# hosts: cluster files [SUCCESS=return] nis
# rpc: files nis
```

8. (可选) 定制 `nfsd` 或 `lockd` 启动选项。

- a. 要定制 `nfsd` 选项，请在每个群集节点上打开 `/etc/init.d/nfs.server` 文件，查找以 `/usr/lib/nfs/nfsd` 开始的命令行，然后添加所需的任何其它变量。

- b. 要定制 `lockd` 启动选项，请在每个群集节点上打开 `/etc/init.d/nfs.client` 文件，查找以 `/usr/lib/nfs/lockd` 开始的命令行，然后添加所需的任何命令行变量。

从 Solaris 9 开始，您可以在 `/etc/default/nfs` 文件中使用 `LOCKD_GRACE_PERIOD` 参数设置 `lockd` 宽延期。但是，如果在 `/etc/init.d/nfs.client` 文件中的命令行变量中设置了宽延期，这将覆盖 `LOCKD_GRACE_PERIOD` 中设置的值。

注意 – 这些命令行必须在一行内。不支持断为多行的命令行。其它变量必须为 `nfsd (1M)` 和 `lockd (1M)` 手册页中介绍的有效选项。

9. 将所需的逻辑主机名资源添加到故障切换资源组中。

您必须按此步骤设置逻辑主机名资源。要用于 Sun Cluster HA for NFS 的逻辑主机名不能为 `SharedAddress` 资源类型。

```
# scrgadm -a -L -g resource-group -l logical-hostname, ... [-n netiflist]
-a                指定您要添加新配置。
-L -g resource-group 指定用于保存逻辑主机名资源的资源组。
-l logical-hostname, ... 指定要添加的逻辑主机名资源。
-n netiflist      指定一个可选的、用逗号分隔的列表，该列表标识各节点上的 IP 网络多路径处理组。netiflist 中的每个元素的格式必须为 netif@node。netif 可被指定为 IP 网络多路径
```


处理组的名称，例如 `sc_ipmp0`。节点可由节点名称或节点 ID 标识，例如 `sc_ipmp0@1` 或 `sc_ipmp@phys-schost-1`。

注意 – Sun Cluster 目前不支持将适配器名称用作 `netif`。

10. 在任意群集节点上为 NFS 配置文件创建一个目录结构。

在 `Pathprefix` 特性于步骤 4 中标识的目录下面创建管理子目录，例如 `/global/nfs/SUNW.nfs`。

```
# mkdir Pathprefix/SUNW.nfs
```

11. 在步骤 10 中创建的 `SUNW.nfs` 目录中创建一个 `dfstab.resource` 文件，并设置共享选项。

a. 创建 `Pathprefix/SUNW.nfs/dfstab.resource` 文件。

此文件包含一组使用共享路径名的 `share` 命令。共享路径应为群集文件系统的子目录。

注意 – 选择一个 `resource` 名称后缀以标识要创建的（在步骤 13 中）NFS 资源。一个好的资源名称可以指代此资源要执行的任务。例如，对于共享用户起始目录的 NFS 资源而言，名称 `user-nfs-home` 就是一个好的候选资源名称。

b. 为您创建的、要共享的每个路径设置 `share` 选项。

此文件的格式与 `/etc/dfs/dfstab` 文件中所用的格式完全相同。

```
share [-F nfs] [-o] specific_options [-d "description"] pathname
```

`-F nfs` 将文件系统类型标识为 `nfs`。

`-o specific_options` 授予所有客户机读写访问权限。有关选项的列表，请参见 `share(1M)` 手册页。为 Sun Cluster 设置 `rw` 选项。

`-d description` 说明要添加的文件系统。

`pathname` 标识要共享的文件系统。

设置共享选项时，请注意以下几点。

- 构造 `share` 选项时，请不要使用 `root` 选项，并且不要将 `ro` 和 `rw` 选项同时使用。
- 不要授权访问群集互连上的主机名。

将读写访问权限授予所有群集节点和逻辑主机，以使 Sun Cluster HA for NFS 监视功能可以执行完整的任务。但是，您可以限制对文件系统的写入访问，也可以使整个文件系统处于只读状态。如果执行了上述操作，Sun Cluster HA for NFS 故障监视功能仍可以在不具备写入权限的情况下执行监视。

- 如果在 `share` 命令中指定一个客户机列表，请包含与群集相关的所有物理主机名和逻辑主机名以及群集连接的所有公共网络中的所有客户机的主机名。
- 如果在 `share` 命令中使用网络组（而不是单个主机的名称），请将所有此类群集主机名添加到相应的网络组。

`share -o rw` 命令将授予所有客户机（包括 Sun Cluster 软件使用的主机名）写入权限。此命令能够使 Sun Cluster HA for NFS 故障监视功能的运行效率最佳。有关详细信息，请参见以下手册页。

- `dfstab(4)`
- `share(1M)`
- `share_nfs(1M)`

12. 注册 NFS 资源类型。

```
# scrgadm -a -t resource-type
```

`-a -t resource-type` 添加指定的资源类型。对于 Sun Cluster HA for NFS，资源类型为 `SUNW.nfs`。

13. 在故障切换资源组中创建 NFS 资源。

```
# scrgadm -a -j resource -g resource-group -t resource-type
```

`-a` 指定您要添加配置。

`-j resource` 指定要添加的、在步骤 11 中定义的资源名称。您可以选择此名称，但它在群集中必须唯一。

`-g resource-group` 指定先前创建的、要添加此资源的资源组的名称。

`-t resource-type` 指定此资源所属的资源类型的名称。此名称必须是已注册的资源类型的名称。

14. 运行 `scswitch(1M)` 命令以执行以下任务。

- 启用资源和资源故障监视器。
- 管理资源组。
- 将资源组切换到 ONLINE 状态。

```
# scswitch -Z -g resource-group
```

示例——设置和配置 Sun Cluster HA for NFS

下例说明了如何设置和配置 Sun Cluster HA for NFS。

```

( 创建逻辑主机资源组并指定 NFS 所用的管理文件的路径(Pathprefix)。 )
# scrgadm -a -g resource-group-1 -y Pathprefix=/global/nfs

( 将逻辑主机名资源添加到逻辑主机资源组中。 )
# scrgadm -a -L -g resource-group-1 -l schost-1

( 使目录结构包含 Sun Cluster HA for NFS 配置文件。 )
# mkdir -p /global/nfs/SUNW.nfs

( 在 nfs/SUNW.nfs 目录下创建 dfstab.resource 文件并设置共享选项。 )
# share -F nfs -o rw=engineering -d "home dirs" nfs/SUNW.nfs

( 注册 NFS 资源类型。 )
# scrgadm -a -t SUNW.nfs

( 在资源组中创建 NFS 资源。 )
# sc
rgadm -a -j r-nfs -g resource-group-1 -t SUNW.nfs

( 启用资源及其监视器，管理资源组并将资源组切换到联机状态。 )
# scswitch -Z -g resource-group-1

```

▼ 如何更改 NFS 文件系统上的共享选项

如果使用 `share -o` 命令的 `rw`、`rw=`、`ro` 或 `ro=` 选项，则 NFS 故障监视功能在以下情况下运行效果最佳：授权访问与所有 Sun Cluster 服务器相关的所有物理主机或 `netgroups`。

如果在 `share(1M)` 命令中使用 `netgroups`，请将所有 Sun Cluster 主机名添加到相应的 `netgroup` 中。理想的做法是将读写访问权限授予所有 Sun Cluster 主机名，以使 NFS 故障探测可以执行完整的任务。

注意 – 更改共享选项之前，请阅读 `share_nfs(1M)` 手册页以了解哪些选项组合是有效的。

您还可以动态修改共享路径和选项而无需使 Sun Cluster HA for NFS 资源脱机。请参见第 21 页“如何动态更新 NFS 文件系统上的共享路径”。

要在 Sun Cluster HA for NFS 资源处于脱机状态时修改 NFS 文件系统上的共享选项，请执行以下步骤。

1. 成为群集节点的超级用户。
2. 关闭 NFS 资源的故障监视功能。

```
# scswitch -n -M -j resource
```

-M 禁用资源故障监视器

3. 测试新的 `share` 选项。

- a. 在编辑带有新的共享选项的 `dfstab.resource` 文件之前，请执行新的 `share` 命令以检验选项组合的有效性。

```
# share -F nfs [-o] specific_options [-d "description"] pathname
-F nfs          将文件系统类型标识为 NFS。
-o specific_options 指定一个选项。您可以使用 rw，它可将读写访问权限授予所有客户机。
-d description  说明要添加的文件系统。
pathname       标识要共享的文件系统。
```

- b. 如果新的 `share` 命令失败，请立即执行另一个带有旧选项的 `share` 命令。如果新的命令成功执行，请继续步骤 4。

4. 编辑带有新的共享选项的 `dfstab.resource` 文件。

- a. 要删除 `dfstab.resource` 文件中的路径，请按顺序执行以下步骤。

- i. 执行 `unshare(1M)` 命令。
- ii. 在 `dfstab.resource` 文件中，删除要删除路径的 `share` 命令。

```
# unshare [-F nfs] [-o specific_options] pathname
# vi dfstab.resource
-F nfs          将文件系统类型标识为 NFS。
-o             指定特定于 NFS 文件系统的选项。
specific_options
pathname       标识不可用的文件系统。
```

- b. 要在 `dfstab.resource` 文件中添加路径或更改现有路径，请先检验装载点是否有效，然后编辑 `dfstab.resource` 文件。

注意 - 此文件的格式与 `/etc/dfs/dfstab` 文件中所用的格式完全相同。每行包括一个 `share` 命令。

5. 启用 NFS 资源的故障监视功能。

```
# scswitch -e -M -j resource
```

▼ 如何动态更新 NFS 文件系统上的共享路径

您可以动态修改 NFS 文件系统上的共享路径而无需使 Sun Cluster HA for NFS 资源脱机。一般步骤包括：修改 Sun Cluster HA for NFS 的 `dfstab.resource` 文件，然后手动运行相应的命令（`share` 或 `unshare` 命令）。命令将立即生效，而 Sun Cluster HA for NFS 将使这些路径具有高可用性。

确保在故障转移期间 Sun Cluster HA for NFS 始终可以使用这些共享路径，从而确保不使用本地路径（在非 HA 文件系统中）。

如果共享了 HAStoragePlus 所管理的文件系统上的路径，则 HAStoragePlus 资源必须与 Sun Cluster HA for NFS 资源在同一资源组中，并且必须正确设置这些资源之间的依赖性。

1. 使用 `scstat -g` 命令查找 Sun Cluster HA for NFS 资源处于联机状态的节点。
2. 在此节点上运行 `/usr/sbin/share` 命令查看当前共享的路径的列表。确定要对此列表进行的更改。
3. 要添加新的共享路径，请执行以下步骤。
 - a. 将 `share` 命令添加到 `dfstab.resource` 文件中。

Sun Cluster HA for NFS 将在下次检查该文件时共享新的路径。检查的频率由 `Thorough_Probe_Interval` 特性来控制（缺省值为 120 秒）。
 - b. 手动运行 `share` 命令使新添加的共享路径立即生效。建议手动运行此命令，因为这样做用户可确保潜在客户机能够使用共享路径。Sun Cluster HA for NFS 将检测是否已共享新添加的路径且未报错。
4. 要取消路径共享，请执行以下步骤。
 - a. 运行 `dfmounts (1M)` 命令以确保当前没有客户机使用此路径。

虽然可在客户机仍在使用的路径时取消路径共享，但这样做客户机将得到无效文件错误句柄并需要执行特殊操作（强制卸载，甚至重新引导）才能恢复。
 - b. 删除 `dfstab.resource` 文件中的共享路径。
 - c. 手动运行 `unshare` 命令。
5. 要修改现有共享路径的选项，请执行以下步骤。
 - a. 根据需要修改 `dfstab.resource` 文件。
 - b. 手动运行相应的命令（`share` 或 `unshare`）。

如何调整 Sun Cluster HA for NFS 方法超时

Sun Cluster HA for NFS 方法完成所需要的时间取决于资源通过 `dfstab.resource` 文件共享的路径的数目。这些方法的缺省超时值为 300 秒。

一般情况下，为每个共享路径分配 10 秒的方法超时值。缺省超时值设计为处理 30 个共享路径。

- 如果共享路径的数目小于 30，请勿减少超时值。
- 如果共享路径的数目大于 30，请将路径数目乘以 10 来计算建议的超时值。例如，如果 `dfstab.resource` 文件包含 50 个共享路径，建议的超时值为 500 秒。

如果共享路径的数目大于 30，请更新以下方法超时。

<code>Prenet_start_timeout</code>	<code>Postnet_stop_timeout</code>	<code>Monitor_Start_timeout</code>
<code>Start_timeout</code>	<code>Validate_timeout</code>	<code>Monitor_Stop_timeout</code>
<code>Stop_timeout</code>	<code>Update_timeout</code>	<code>Monitor_Check_timeout</code>

要更改方法超时，请使用 `scrgadm -c` 选项（如下例所示）。

```
% scrgadm -c -j resource -y Prenet_start_timeout=500
```

如何配置 SUNW.HAStoragePlus 资源类型

Sun Cluster 3.0 5/02 中新增了 SUNW.HAStoragePlus 资源类型。这种新资源类型可以执行与 SUNW.HAStorage 相同的功能，并可以同步 HA 存储和 Sun Cluster HA for NFS 之间的操作。SUNW.HAStoragePlus 还具备使本地文件系统具有高可用性的附加功能。Sun Cluster HA for NFS 是磁盘密集型程序，并且可以进行故障切换，因此您应当设置 SUNW.HAStoragePlus 资源类型。

有关背景信息，请参阅 SUNW.HAStoragePlus (5) 手册页和《Sun Cluster 数据服务和规划指南（适用于 Solaris OS）》中的“Relationship Between Resource Groups and Disk Device Groups”。有关过程信息，请参见《用于 Solaris OS 的 Sun Cluster 数据规划和管理指南》中的“在资源组和磁盘设备组之间同步启动”。（如果您使用的是 5/02 之前的 Sun Cluster 3.0 版本，则必须设置 SUNW.HAStorage 而不是 SUNW.HAStoragePlus。有关过程信息，请参见《用于 Solaris OS 的 Sun Cluster 数据服务规划和管理指南》中的“在资源组和磁盘设备组之间同步启动”。）

配置 Sun Cluster HA for NFS 扩展特性

通常，在创建 NFS 资源时使用命令行 `scrgadm -x parameter =value` 来配置扩展特性。您以后还可以使用《Sun Cluster 数据服务规划和管理指南（适用于 Solaris OS）》中的“Administering Data Service Resources”中的过程来配置这些特性。您无需为 Sun Cluster HA for NFS 设置任何扩展特性。有关所有 Sun Cluster 特性的详细信息，请参见《Sun Cluster 数据服务规划和管理指南（适用于 Solaris OS）》中的“标准特性”。

有关可以为 Sun Cluster HA for NFS 配置的扩展特性的信息，请参阅表 2。您可以动态地更新某些特性。但是，您只有在创建资源时才可以更新其它特性。“可调”条目说明何时可以更新该特性。

表 2 Sun Cluster HA for NFS 扩展特性

名称/数据类型	缺省值
Lockd_nullrpc_timeout (整数)	探测 lockd 时所用的超时值 (以秒为单位)。 缺省值 : 120 范围 : 最小值 = 60 可调 : 随时
Monitor_retry_count (整数)	在 Monitor_retry_interval 特性指定的时间窗口中进程监视工具 (PMF) 重新启动故障监视器的次数。请注意，此特性指的是重新启动故障监视器本身 (而不是资源)。系统定义的特性 Retry_interval 和 Retry_count 用于控制资源的重新启动。有关这些特性的说明，请参见 scrgadm(1M) 手册页。 缺省值 : 4 范围 : 0 - 2,147,483,641 -1 表示重新启动无限次。 可调 : 随时
Monitor_retry_interval (整数)	记录故障监视器失败次数的时间段 (以分钟为单位)。如果故障监视器的失败次数超过了扩展特性 Monitor_retry_count 在此时间段中指定的值，则 PMF 将重新启动故障监视器。 缺省值 : 2 范围 : 0 - 2,147,483,641 -1 表示时间无限长。 可调 : 随时
Mountd_nullrpc_restart (布尔值)	布尔值，表示在空 rpc 调用失败时是否重新启动 mountd。 缺省值 : True 范围 : 无 可调 : 随时

表 2 Sun Cluster HA for NFS 扩展特性 (续)

名称/数据类型	缺省值
Mountd_nullrpc_timeout (整数)	探测 mountd 时所用的超时值 (以秒为单位)。 缺省值: 120 范围: 最小值 = 60 可调: 随时
Nfsd_nullrpc_restart (布尔值)	布尔值, 表示在空 rpc 调用失败时是否重新启动 nfsd。 缺省值: False 范围: 无 可调: 随时
Nfsd_nullrpc_timeout (整数)	探测 nfsd 时所用的超时值 (以秒为单位)。 缺省值: 120 范围: 最小值 = 60 可调: 随时
Rpcbind_nullrpc_reboot (布尔值)	布尔值, 表示针对 rpcbind 的空 rpc 调用失败时是否重新引导系统。 缺省值: False 范围: 无 可调: 随时
Rpcbind_nullrpc_timeout (整数)	探测 rpcbind 时所用的超时值 (以秒为单位)。 缺省值: 120 范围: 最小值 = 60 可调: 随时
Statd_nullrpc_timeout (整数)	探测 statd 时所用的超时值 (以秒为单位)。 缺省值: 120 范围: 最小值 = 60 可调: 随时

Sun Cluster HA for NFS 故障监视器

Sun Cluster HA for NFS 故障监视器使用以下两个进程：

- NFS 系统故障监视，该进程包括监视 NFS 守护程序（`nfsd`、`mountd`、`statd` 和 `lockd`）和解决发生的所有问题。NFS 系统故障监视进程还监视 RPC 端口映射器服务守护程序（`rpcbind`）。
- NFS 资源故障监视，该进程针对每个 NFS 资源。每个资源的故障监视器检查每个共享路径的状态，以监视资源输出的文件系统。

启动故障监视器

首先，NFS 资源 `MONITOR_START` 方法启动 NFS 系统故障监视器。此启动方法先检查 NFS 系统故障监视器（`nfs_daemons_probe`）是否已在进程监视器守护程序（`rpc.pmfd`）下运行。如果 NFS 系统故障监视器未运行，启动方法将在进程监视器的控制下启动 `nfs_daemons_probe` 进程。然后，启动方法将仍然在进程监视器的控制下启动资源故障监视器（`nfs_probe`）。

停止故障监视器

首先，NFS 资源 `MONITOR_STOP` 方法停止资源故障监视器。然后，如果本地节点上没有运行其它 NFS 资源故障监视器，此方法将停止 NFS 系统故障监视器。

NFS 系统故障监视器进程

NFS 系统故障监视器通过检查进程是否存在及其对空 `rpc` 调用的响应，来探测本地节点上的 `rpcbind`、`statd`、`lockd`、`nfsd` 和 `mountd`。此监视器使用以下 NFS 扩展特性。

<code>Rpcbind_nullrpc_timeout</code>	<code>Lockd_nullrpc_timeout</code>
<code>Nfsd_nullrpc_timeout</code>	<code>Rpcbind_nullrpc_reboot</code>
<code>Mountd_nullrpc_timeout</code>	<code>Nfsd_nullrpc_restart</code>
<code>Statd_nullrpc_timeout</code>	<code>Mountd_nullrpc_restart</code>

请参见第 22 页“配置 Sun Cluster HA for NFS 扩展特性”以查看或设置扩展特性。

每个系统故障监视器探测周期将在一个循环中执行以下步骤。

1. 进行休眠，时间为 `Cheap_probe_interval`。

2. 探测 `rpcbind`。

如果进程意外终止，但是守护程序正在进行重新热启动，`rpcbind` 将继续探测其它守护程序。

如果进程意外终止，则故障监视器将重新引导节点。

如果针对守护程序的空 `rpc` 调用意外终止，并且

`Rpcbind_nullrpc_reboot=True`，`Failover_mode=HARD`，则故障监视器将重新引导节点。

3. 先探测 `statd`，然后探测 `lockd`。

如果 `statd` 或 `lockd` 意外终止，则系统故障监视器将尝试重新启动这两个守护程序。

如果针对这些守护程序的空 `rpc` 调用意外终止，则故障监视器将在 `syslog` 中记录一条消息，但是不会重新启动 `statd` 或 `lockd`。

4. 探测 `mountd`。

如果 `mountd` 意外终止，则故障监视器将尝试重新启动该守护程序。

如果针对守护程序的空 `rpc` 调用意外终止，并且

`Mountd_nullrpc_restart=True`，则故障监视器将在群集文件系统可用时尝试重新启动 `mountd`。

5. 探测 `nfsd`。

如果 `nfsd` 意外终止，则故障监视器将尝试重新启动该守护程序。

如果针对守护程序的空 `rpc` 调用意外终止，并且

`Nfsd_nullrpc_restart=TRUE`，则故障监视器将在群集文件系统可用时尝试重新启动 `nfsd`。

如果上述任何 NFS 守护程序（`rpcbind` 除外）在探测周期期间无法重新启动，则 NFS 系统故障监视器将在下一个周期内再次尝试重新启动。当所有 NFS 守护程序均已重新启动并且运行正常时，资源状态将被设置为 `ONLINE`。监视器在上一个 `Retry_interval` 内记录 NFS 守护程序的意外终止。如果守护程序的意外终止总数达到了 `Retry_count`，则系统故障监视器将发出 `scha_control` 停止通知。如果停止调用失败，监视器将尝试重新启动失败的 NFS 守护程序。

在每个探测周期的末尾，如果所有守护程序均正常运行，监视器将清除故障历史记录。

NFS 资源故障监视器进程

启动资源故障监视器探测之前，请从 `dfstab` 文件读取所有共享路径并将它们存储在内存中。在每个探测周期内，通过对路径执行 `stat()` 以探测每个迭代中的所有共享路径。

每个资源故障监视器故障探测将在一个循环中执行以下步骤。

1. 进行休眠，时间为 `Thorough_probe_interval`。

2. 如果 `dfstab` 自上次读取后已发生更改，则刷新内存。
如果在读取 `dfstab` 文件时发生错误，则资源状态将被设置为 `FAULTED`，监视器将跳过当前探测周期的剩余检查。
3. 通过对路径执行 `stat()` 以探测每个迭代中的所有共享路径。
如果任何路径不能正常运行，则资源状态将被设置为 `FAULTED`。
4. 探测 NFS 守护程序 (`nfsd`、`mountd`、`lockd` 和 `statd`) 以及 `rpcbind` 是否存在。
5. 如果这些守护程序中的任何一个停止，则资源状态将被设置为 `FAULTED`。

如果所有共享路径都有效并且 NFS 守护程序都存在，则资源状态将在探测周期的末尾被重置为 `ONLINE`。

索引

C

C 语言环境, 13

P

prtconf -v 命令, 9

prtdiag -v 命令, 9

psrinfo -v 命令, 9

S

scinstall -pv 命令, 9

showrev -p 命令, 9

Sun Cluster HA for Network File System (NFS)

安装

通过使用 Web Start 程序, 13-14

Sun Cluster HA for NFS

SUNW.HAStoragePlus 资源类型, 22

安装

安装软件, 14

调整方法超时, 21

更改共享选项, 19, 21

故障监视器, 25

扩展特性

Lockd_nullrpc_timeout, 22

Monitor_retry_count, 22

Monitor_retry_interval, 22

Mountd_nullrpc_restart, 22

Mountd_nullrpc_timeout, 22

Nfsd_nullrpc_restart, 22

Nfsd_nullrpc_timeout, 22

Sun Cluster HA for NFS, 扩展特性 (续)

Rpcbind_nullrpc_reboot, 22

Rpcbind_nullrpc_timeout, 22

Statd_nullrpc_timeout, 22

任务对应关系, 12

注册和配置, 15

SUNW.HAStoragePlus 资源类型, Sun Cluster

HA for NFS, 22

V

/var/sadm/install/logs 目录, 13

W

Web Start 程序, 13-14

安

安装

Sun Cluster HA for Network File System (NFS)

创建的日志文件, 13

通过使用 Web Start 程序, 13-14

Sun Cluster HA for NFS, 14

方

方法超时, Sun Cluster HA for NFS, 21

共

共享选项

Sun Cluster HA for NFS, 19, 21

故

故障监视器

Sun Cluster HA for NFS, 25

进程, 25

启动, 25

停止, 25

资源故障监视器进程, 26

扩

扩展特性

Sun Cluster HA for NFS

Lockd_nullrpc_timeout, 22

Monitor_retry_count, 22

Monitor_retry_interval, 22

Mountd_nullrpc_restart, 22

Mountd_nullrpc_timeout, 22

Nfsd_nullrpc_restart, 22

Nfsd_nullrpc_timeout, 22

Rpcbind_nullrpc_reboot, 22

Rpcbind_nullrpc_timeout, 22

Statd_nullrpc_timeout, 22

命

命令, 节点信息, 8

目

目录, /var/sadm/install/logs, 13

配

配置, Sun Cluster HA for NFS, 15

任

任务对应关系, Sun Cluster HA for NFS, 12

日

日志文件, 安装, 13

文

文件, 安装日志, 13

语

语言环境, 13

注

注册, Sun Cluster HA for NFS, 15